International Journal of Innovative Research in Computer Science & Technology (IJIRCST)
ISSN: 2347-5552, Volume-2, Issue-2, March-2014

Portable, Robust and Effective Text and Product Label Reading, Currency and Obstacle Detection For Blind Persons

Asha Mohandas, Bhagyalakshmi. G, Manimala. G

Abstract- **The proposed system is a camera-based assistive text reading framework to help blind persons detect currency and identify the obstacle in front in addition to read text labels and product packaging from hand-held objects . To identify text from cluttered backdrops, the image is first converted to gray, then to binary form. Text is localized using text localization algorithm and haar cascade is employed to identify the text and then e-speak engine converts the text into voice output. The above system is enhanced to recognize any Obstacle in front and produce the voice output through Ear phone to blind users. Adaboost model is employed in the obstacle detection process wherein it identifies human and vehicles apart and is given as voice output. To develop a novel camera-based computer vision technology to automatically recognize banknotes to assist visually impaired people. A novel component-based banknote recognition system by using SURF (Speeded Up Robust Features) to achieve high recognition accuracy and to handle various challenging conditions in real world environments is proposed. The input image is converted to gray, from which the descriptors and key points of note are extracted and compared with template and the voice output is produced through Earphone to blind users.**

Index Terms: **Identify Banknotes, Obstacle Detection, SURF application, Text localization.**

I. INTRODUCTION

Blind people find very difficult to read the text in the newspapers, labels, reports, receipts, bank statements, restaurant menus, product packages, instructions on medicine bottles etc. There are no proper systems to help the blind people to identify the text, especially product labels. Blind people find it very hard to locate the bar code area in order to find out the product name. Blind people also find difficult to identify the number on the currency notes and identify the obstacle in front in a clear manner such that they don't trip and fall. There is no technique prevailing in this area till date. In this paper, we propose a single utility which helps blind in an absolutely easy manner. This utility consists of a camera and earphone. The USB camera captures the video of the object in front, may it be text labels, currency notes, human, vehicle or animal. The video is processed and the exact voice output is produced. This is indeed a functional eye for any blind person.

Manuscript received March 24, 2014.
Asha Mohandas, Computer Science and Engineering, Sri Sai Ram Engineering College, Chennai, India, 8939760195
(e-mail: ashamohandas@rocketmail.com).
G. Bhagyalakshmi, Computer Science and Engineering, Sri Sai Ram Engineering College, Chennai, India, 9962071712
(e-mail: bhagyalakshmig2011@gmail.com).
G. Manimala, Computer Science and Engineering, Sri Sai Ram Engineering College, Chennai, India, 9884675773
(e-mail: manimala.cse@sairam.edu.in).

II. TEXT AND PRODUCT LABEL READING

The live video is captured by using web cam and it can be done using OPENCV libraries. The image format from the webcam is in RGB24 format. The frame from the video is segregated and undergoes pre processing. The captured videos are projected in a window with a size of 320x240. Totally 10 frames per second can be captured by using the webcam.

To identify the text from the input image, first the video captured is segregated into frames and each frame is converted first into gray image and then into binary image. Then *Text Localization* algorithm is applied on the binary image to localize the text from the background. The output is given to the *Optical Character Recognition (OCR)[2],[3]* for the text recognition and the audio output is generated.

A. Text Localization Algorithm

Text localization algorithm[2],[5] takes the frames which are segregated from the video as the input and the Region Of Interest (ROI) is found by taking only the required text area from the image. The ROI region is confined within the rectangular area contain the text which is to be get detected, and that text which is inside the ROI is first converted to the gray image and then to binary image. The region of interest is specified and the localized text gets compared with the predefined text in the training set by the Adaboost Model. Adaboost model is in charge for identifying which alphabet is exactly located on the identified region.

B. Haar Cascade-Adaboost Model

Adaboost model [1],[3] is the effective machine learning training set for the text detection. The training set consists of positive and negative samples where the positive samples comprise of the text images and the negative samples comprise of the images other than the text present in the input image. The input image is given to Haar Cascade [1],[3] to analyze the text which is inside the ROI and match the input text with the predefined text in the training text. Both the existing training set and the given input image are compared to identify variations. The one to which the range

of similarity is high is then confirmed to be the character and the corresponding audio is produced.

C. Audio Output Using E-Speak Engine

Optical Character Recognition (OCR) [2],[3] is the mechanical or electronic conversion of images of typewritten or printed text into machine-encoded or computer-readable text. The output of the haar cascade is stored in out.png file which dynamically overwrites the text with every frame. The processed output is send to the e-speak engine which is based on Microsoft Speech Synthesizer Development kit and the audio output is generated and given out via ear phone.

D. Block Diagram

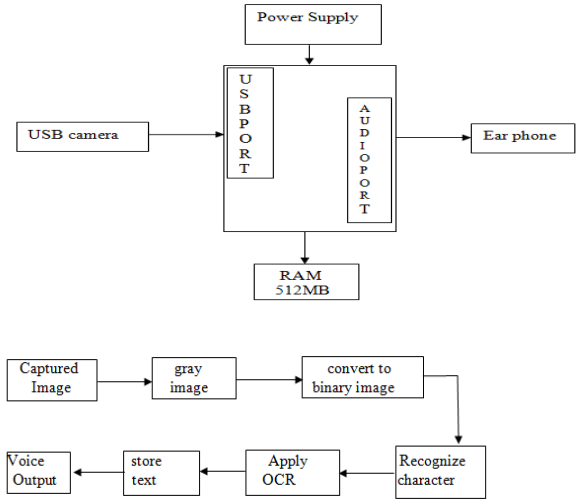

Fig 1: Explains the architecture and components of Text Detection.

E. Flow Chart

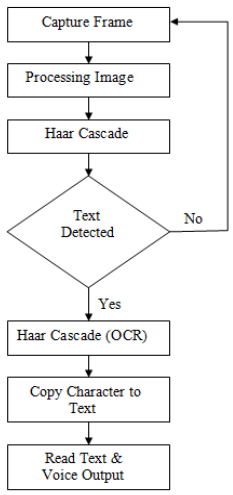

Fig 2: Explains the process flow of Text Detection.

III. CURRENCY NOTE DETECTION

SURF (Speeded Up Robust Features) [6] is a robust local feature detector, to identify the pixel intensities of the images. Feature detection is a low-level image processing operation. That is, it is usually performed as the first operation on an image, and examines every pixel to see if there is a feature present at that pixel. Many computer vision algorithms use feature detection as the initial step, so as a result, a very large number of feature detectors have been developed.

A. Distinguish Interest Point Descriptors

The currency note is laid out in front of the camera and from the captured image SURF *(Speeded Up Robust Features)* [6] detector is used to observe the pixel intensities. Each points in the currency note has different color ranges, the pixel intensities differs for each color range .Those descriptors are defined by a circle. The input image gets reformed into gray image and then reformed into binary image.

B. Distinguish Key Points

Once the *descriptors [6]* are identified, then the corresponding key points are gathered. The input currency note's descriptors and *key points* are gathered and stored separately. Then the predefined currency note's *descriptors* and *key points [6]* are gathered from the data set and stored separately. Each key point has a decimal value and those decimal values are summed up and that threshold value is compared with predefined threshold value. If the summed up decimal values are around 0.5 to 1.0, then the image is matched with the corresponding object. OCR which can handle images with complex background converts the images into text and gives the output to the e-speak engine. The audio is produced and given out via the earphone.

C. Block Diagram

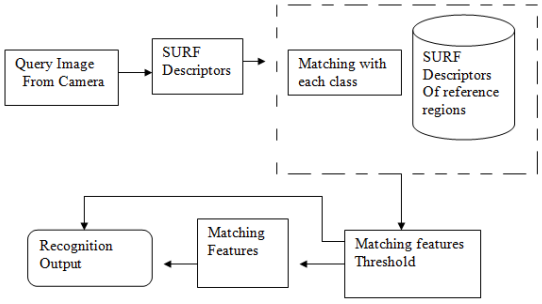

Fig 3: Explains the architecture of Bank Note Detection.

D. Flow Chart

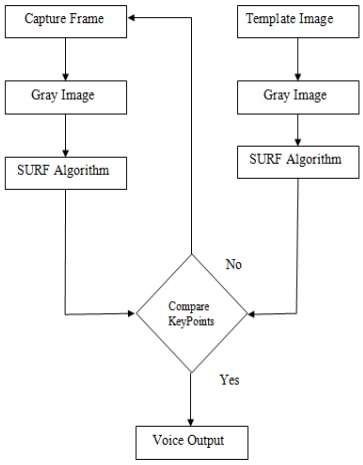

Fig 4: Explains the process flow of Bank Note Detection.

IV. OBSTACLE DETECTION

To identify the type of the obstacle in the input image, the image gets compared with the samples in the training set. Then the positive image alone gets localized from the background by the Adaboost Model .The output is given to the Optical Character Recognition(OCR) [2][3] for the text recognition and the audio output is generated.

A. Cascade- Adaboost Model

Adaboost model [1][3] is the effective machine learning training set for the text detection. The training set consists of positive and negative samples where the positive samples comprise of the human face or car images and the negative samples comprise of the images other than the object (human face or car) present in the input image. The input image is given to Cascade Adaboost Model to analyze the image and match the input image with the predefined image in the training set. Then the corresponding output is produces as audio. Apart from human or car, any object can be included in the training set so as to help the blind person. Our training set is restricted to human and car.

B. Block Diagram

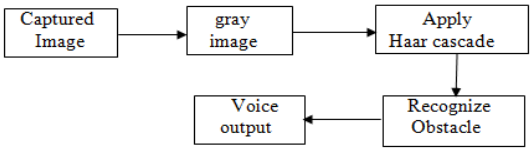

Fig 5: Explains the Components of Obstacle Detection.

C. Flow Chart

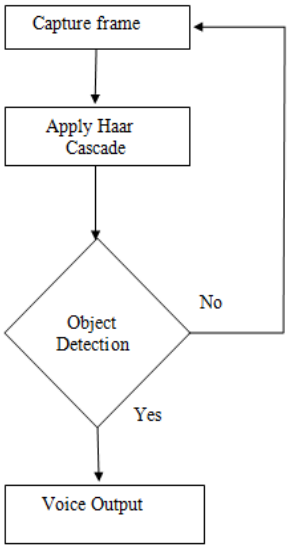

Fig 6: Explains the process flow of Obstacle Detection.

V. RASPBERRY PI

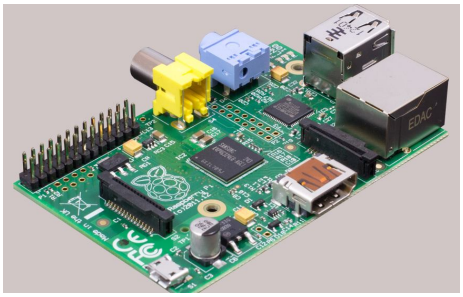

Fig 7(a): Explains the Structure and Components of Raspberry PI Board.

Fig 7 depicts the kit used for the execution is a Raspberry PI board which comprises of various slots for interfacing with the computer, earphone, power supply and the external USB camera. The computer acts as a remote to invoke the program.

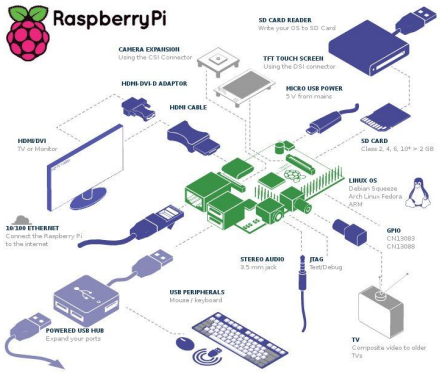

Fig 7(b):The various available slots on a Raspberry PI Board.

VI. PROJECT RESULTS

A. *The Original Setup*

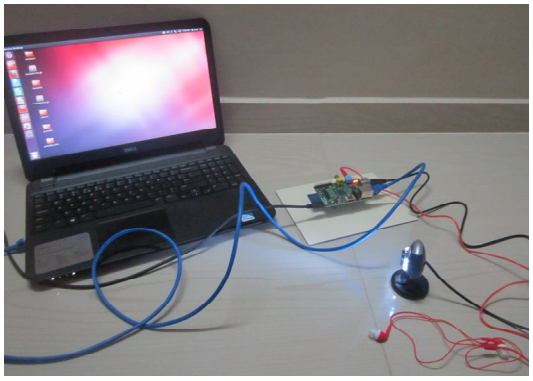

Fig 8: The Total Setup of the entire Kit.

Fig 8 depicts the entire original setup of how the connected system appears. The Laptop acts as a remote to invoke the program. Linux is the platform used where OPEN_CV library is used. The blue wire depicts the LAN interfacing between the OS loaded in the memory card present in the kit and the laptop (Remote).

B. *Connections On The Raspberry PI Board*

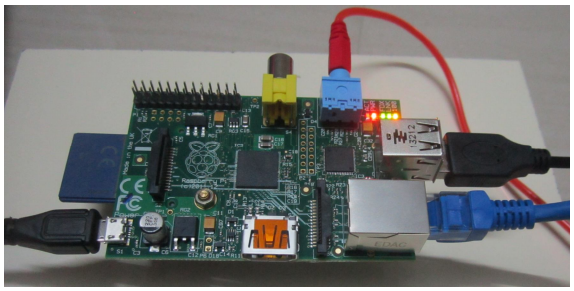

Fig 9: The Connections on the board.

In Fig 9 , the left end consists of two slots:
1) The memory card slot
2) The power supply
The right end consists of two slots:
1) The USB port where the external camera is attached
2) The LAN interfacing with the laptop
The top (Red) wire leads to the earphone.

C. *Module 1: Portable Text and Product Label Reading*

Fig 10: The Execution and result of Text Detection.

D. *Module 2: Currency Note Detection*

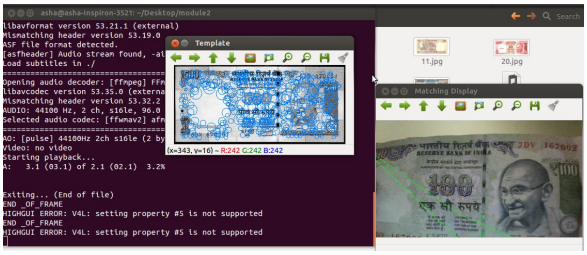

Fig 11: The Execution and result of Bank Note Detection.

E. *Module 3: Obstacle Detection*

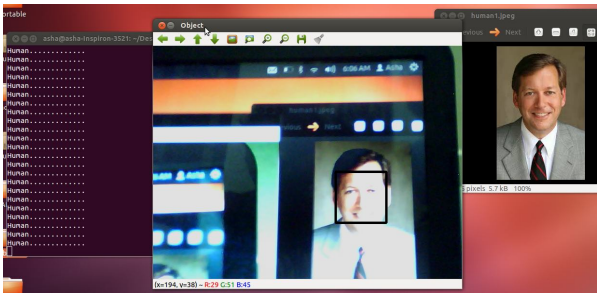

Fig 12(a): The Execution and result of Obstacle Detection-Human.

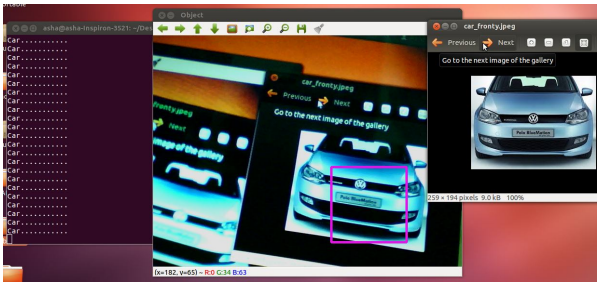

Fig 12(b): The Execution and result of Obstacle Detection-Car.

VII.REFERENCES

[1] Chucai Yi, Student Member, IEEE, Yingli Tian, Senior Member, IEEE, and Aries Arditi, "Portable camera-based assistive text and product label reading from hand-held objects for blind persons",IEEE 2013.
[2] Qixiang Ye , Jianbin Jiao, Jun Huang, Hua Yu, "Text detection and restoration in natural scene images",July 2007.
[3] X. Chen and A. L. Yuille, "Detecting and reading text in natural scenes," in *Proc. Comput. Vision Pattern Recognit*, 2004, vol. 2, pp. II-366–II-373.
[4] Shehzad Muhammad Hanif, Lionel Prevost, "Texture based text detection in natural scene images A help to blind and visually impaired persons" Conference 2007.
[5] Piyanuch Silapachote, Jerod Weinman, Allen Hanson, Richard Weiss†, and Marwan A. Mattar. "Automatic sign detection and recognition in natural scenes"IEEE 2005.
[6] Herbert Bay, Andreas Ess , Tinne Tuytelaars, and Luc Van Gool,"Speeded-up robust features (SURF)"IEEE Sep.2008
[7] X. Chen, J. Yang, J. Zhang, and A. Waibel, "Automatic detection and Recognition of signs from natural scenes," *IEEE Trans. Image processes*, vol. 13, no. 1, pp. 87–99, Jan. 2004.
[8] D. Dakopoulos and N. G. Bourbakis, "Wearable obstacle avoidance electronic Travel aids for blind: A survey," *IEEE Trans. Syst., Man, Cybern.* vol. 40, no. 1, pp. 25–35, Jan. 2010.

VIII. AUTHOR PROFILE

Asha Mohandas is pursuing her Bachelor of Engineering degree in the Department of Computer Science and Engineering at Sri Sai Ram Engineering College, Anna University, Chennai in the year 2014.

G. Bhagyalakshmi is pursuing her Bachelor of Engineering degree in the Department of Computer Science and Engineering at Sri Sai Ram Engineering College, Anna University, Chennai in the year 2014.

G. Manimala is working as an Associate Professor with 14½ years of Teaching Experience in the Department of Computer Science and Engineering at Sri Sai Ram Engineering College, Chennai. She is interested in Computer and Communication area and has presented and published papers in Journal, National and International Conferences. She is lifetime member